IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of:	)	
Inventor: Judith A. Bayer et al.	)	Examiner: Daniel Lastra
Serial #: 09/998,750	)	Group Art Unit: 3622
Filed: November 30, 2001	)	Appeal No.: _________
Title: AUTOMATED PROMOTION RESPONSE MODELING IN A CUSTOMER RELATIONSHIP MANAGEMENT SYSTEM	)	

BRIEF OF APPELLANTS

MAIL STOP APPEAL BRIEF - PATENTS
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In accordance with 37 CFR §41.37, Appellants' attorney hereby submits the Brief of Appellants on appeal from the final rejection in the above-identified application, as set forth in the Office Action dated March 13, 2009.

The Office is authorized to charge any necessary fees or credit any overpayments to Deposit Account No. 50-4370 of Teradata Corporation, the assignee of the present application.

I. REAL PARTY IN INTEREST

The real party in interest is Teradata Corporation, the assignee of the present application.

II. RELATED APPEALS AND INTERFERENCES

There are no related appeals or interferences for the above-referenced patent application.

III. STATUS OF CLAIMS

Claims 1-21 are pending in the application.

Claims 1-21 were rejected under 35 U.S.C. §102(e) as being anticipated by Cook, U.S. Patent No. 6,631,360 (Cook).

Claims 1-21 are being appealed.

IV. STATUS OF AMENDMENTS

No amendments have been submitted subsequent to the final Office Action dated March 13, 2009.

V. SUMMARY OF CLAIMED SUBJECT MATTER

The claimed subject matter can be found in the Appellants' specification as filed as shown below:

Claim Element	**Support in Specification**
1. A computer-implemented method of creating a customer promotion response model for use in customer relationship marketing, comprising:	Page 3, lines 6-17; Page 5, lines 18-30 referring to 102, 104, 106 in FIG. 1; and Page 7, lines 26-31 referring to FIG. 3.
(a) generating, in a computer, an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions, wherein the primitives and conditions determine how the Analytic Variables are derived from operational data to produce the input data set, and wherein the Analytic Variables are subdivided into independent variables and their related dependent variables;	Page 6, line 1 through page 7, line 24 referring to 200, 202, 204 in FIG. 2; and Page 9, lines 4-9 referring to 300 in FIG. 3.

Claim Element	**Support in Specification**
(b) splitting, in the computer, the input data set into a test sample and a validation sample;	Page 9, lines 10-13 referring to 302 in FIG. 3.
(c) identifying, in the computer, the independent variables and their related dependent variables using the test sample;	Page 9, lines 14-23 referring to 304 in FIG. 3.
(d) identifying, in the computer, a Transformation Type for each of the identified independent variables and their related dependent variables;	Page 9, lines 24-29 referring to 306 in FIG. 3.
(e) estimating, in the computer, a Coefficient for each of the identified independent variables and their related dependent variables;	Page 10, lines 2-9 referring to 308 in FIG. 3.
(f) generating, in the computer, a Model Equation for each of the identified independent variables and their related dependent variables using the identified Transformation Type and estimated Coefficient;	Page 10, lines 10-16 referring to 310 in FIG. 3.
(g) validating, in the computer, the generated Model Equation by applying it to the validation sample; and	Page 10, lines 17-21 referring to 312 in FIG. 3.
(h) scoring, in the computer, customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing.	Page 11, lines 1-11 referring to 314 in FIG. 3.
2. The method of claim 1, wherein the	Page 9, lines 24-29 referring to 306 in FIG.

Claim Element	Support in Specification
Transformation Type is a mathematical operation that identifies an association between the identified related independent and dependent variables.	3.
3. The method of claim 1, wherein the Coefficient is a relative measure of the identified related independent and dependent variables' contributions to a likelihood of response.	Page 10, lines 2-9 referring to 308 in FIG. 3.
4. The method of claim 1, wherein the Coefficient's sign indicates whether the independent variable is positively or negatively correlated with the dependent variable.	Page 10, lines 2-9 referring to 308 in FIG. 3.
5. The method of claim 1, wherein the Model Equation is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders.	Page 10, lines 10-16 referring to 310 in FIG. 3.
6. The method of claim 1, wherein the validating step (g) further comprises applying the generated Model Equation to the validation sample in order to predict a likelihood of response as compared to an actual response in the validation sample.	Page 10, lines 17-21 referring to 312 in FIG. 3.

Claim Element	**Support in Specification**
7. The method of claim 1, wherein the scoring step (h) further comprises applying the validated Model Equation to the customers retrieved from the database in order to predict responses from the customers in a future promotional campaign.	Page 11, lines 1-11 referring to 314 in FIG. 3.
8. A computer-implemented system for creating a customer promotion response model for use in customer relationship marketing, comprising:	Page 3, lines 6-17; Page 4, lines 6-16 referring to 100, 102, 104, 106 in FIG. 1; Page 5, lines 18-30 referring to 102, 104, 106 in FIG. 1; and Page 7, lines 26-31 referring to FIG. 3.
(a) a computer;	Page 4, lines 6-16 referring to 100, 102, 104, 106 in FIG. 1;
(b) a customer relationship marketing system, performed by the computer, for:	Page 4, lines 6-16 referring to 100, 102, 104, 106 in FIG. 1;
(1) generating an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions, wherein the primitives and conditions determine how the Analytic Variables are derived from operational data to produce the input data set, and wherein the Analytic Variables are subdivided into independent	Page 6, line 1 through page 7, line 24 referring to 200, 202, 204 in FIG. 2; and Page 9, lines 4-9 referring to 300 in FIG. 3.

Claim Element	Support in Specification
variables and their related dependent variables;	
(2) splitting the input data set into a test sample and a validation sample;	Page 9, lines 10-13 referring to 302 in FIG. 3.
(3) identifying the independent variables and their related dependent variables using the test sample;	Page 9, lines 14-23 referring to 304 in FIG. 3.
(4) identifying a Transformation Type for each of the identified independent variables and their related dependent variables;	Page 9, lines 24-29 referring to 306 in FIG. 3.
(5) estimating a Coefficient for each of the identified independent variables and their related dependent variables;	Page 10, lines 2-9 referring to 308 in FIG. 3.
(6) generating a Model Equation for each of the identified independent variables and their related dependent variables using the identified Transformation Type and estimated Coefficient;	Page 10, lines 10-16 referring to 310 in FIG. 3.
(7) validating the generated Model Equation by applying it to the validation sample; and	Page 10, lines 17-21 referring to 312 in FIG. 3.
(8) scoring customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing.	Page 11, lines 1-11 referring to 314 in FIG. 3.
9. The system of claim 8, wherein the Transformation Type is a mathematical operation that identifies an association between the identified related independent and dependent variables.	Page 9, lines 24-29 referring to 306 in FIG. 3.

Claim Element	Support in Specification
10. The system of claim 8, wherein the Coefficient is a relative measure of the identified related independent and dependent variables' contributions to a likelihood of response.	Page 10, lines 2-9 referring to 308 in FIG. 3.
11. The system of claim 8, wherein the Coefficient's sign indicates whether the independent variable is positively or negatively correlated with the dependent variable.	Page 10, lines 2-9 referring to 308 in FIG. 3.
12. The system of claim 8, wherein the Model Equation is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders.	Page 10, lines 10-16 referring to 310 in FIG. 3.
13. The system of claim 8, wherein the logic for validating (7) further comprises logic for applying the generated Model Equation to the validation sample in order to predict a likelihood of response as compared to an actual response in the validation sample.	Page 10, lines 17-21 referring to 312 in FIG. 3.
14. The system of claim 8, wherein the logic for scoring (8) further comprises logic for applying the validated Model Equation to the	Page 11, lines 1-11 referring to 314 in FIG. 3.

Claim Element	Support in Specification
customers retrieved from the database in order to predict responses from the customers in a future promotional campaign.	
15. An article of manufacture comprising a storage device embodying instructions that, when read and executed by a computer, result in the computer performing a method for creating a customer promotion response model for use in customer relationship marketing, comprising:	Page 3, lines 6-17; Page 5, lines 6-30 referring to 102, 104, 106 in FIG. 1; and Page 7, lines 26-31 referring to FIG. 3.
(a) generating, in a computer, an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions, wherein the primitives and conditions determine how the Analytic Variables are derived from operational data to produce the input data set, and wherein the Analytic Variables are subdivided into independent variables and their related dependent variables;	Page 6, line 1 through page 7, line 24 referring to 200, 202, 204 in FIG. 2; and Page 9, lines 4-9 referring to 300 in FIG. 3.
(b) splitting, in the computer, the input data set into a test sample and a validation sample;	Page 9, lines 10-13 referring to 302 in FIG. 3.
(c) identifying, in the computer, the independent variables and their related	Page 9, lines 14-23 referring to 304 in FIG. 3.

Claim Element	**Support in Specification**
dependent variables using the test sample;	
(d) identifying, in the computer, a Transformation Type for each of the identified independent variables and their related dependent variables;	Page 9, lines 24-29 referring to 306 in FIG. 3.
(e) estimating, in the computer, a Coefficient for each of the identified independent variables and their related dependent variables;	Page 10, lines 2-9 referring to 308 in FIG. 3.
(f) generating, in the computer, a Model Equation for each of the identified independent variables and their related dependent variables using the identified Transformation Type and estimated Coefficient;	Page 10, lines 10-16 referring to 310 in FIG. 3.
(g) validating, in the computer, the generated Model Equation by applying it to the validation sample; and	Page 10, lines 17-21 referring to 312 in FIG. 3.
(h) scoring, in the computer, customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing.	Page 11, lines 1-11 referring to 314 in FIG. 3.
16. The article of manufacture of claim 15, wherein the Transformation Type is a mathematical operation that identifies an association between the identified related independent and dependent variables.	Page 9, lines 24-29 referring to 306 in FIG. 3.

Claim Element	Support in Specification
17. The article of manufacture of claim 15, wherein the Coefficient is a relative measure of the identified related independent and dependent variables' contributions to a likelihood of response.	Page 10, lines 2-9 referring to 308 in FIG. 3.
18. The article of manufacture of claim 15, wherein the Coefficient's sign indicates whether the independent variable is positively or negatively correlated with the dependent variable.	Page 10, lines 2-9 referring to 308 in FIG. 3.
19. The article of manufacture of claim 15, wherein the Model Equation is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders.	Page 10, lines 10-16 referring to 310 in FIG. 3.
20. The article of manufacture of claim 15, wherein the validating step (g) further comprises applying the generated Model Equation to the validation sample in order to predict a likelihood of response as compared to an actual response in the validation sample.	Page 10, lines 17-21 referring to 312 in FIG. 3.
21. The article of manufacture of claim 15, wherein the scoring step (h) further	Page 11, lines 1-11 referring to 314 in FIG. 3.

Claim Element	Support in Specification
comprises applying the validated Model Equation to the customers retrieved from the database in order to predict responses from the customers in a future promotional campaign.	

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

1. Claims 1-21 stand rejected under 35 U.S.C. §102(e) as being anticipated by Cook, U.S. Patent No. 6,631,360 (Cook).

VII. ARGUMENT

A. Arguments directed to the first grounds for rejection: Claims 1-21 stand rejected under 35 U.S.C. §102(e) as being anticipated by Cook, U.S. Patent No. 6,631,360 (Cook).

1. Independent claims 1, 8 and 15

On pages (3)-(10) of the Office Action, claims 1-21 were rejected under 35 U.S.C. §102(e) as being anticipated by Cook, U.S. Patent No. 6,631,360 (Cook).

Appellants' attorney respectfully traverses these rejections.

Nonetheless, the Office Action asserts the following:

> Claims 1-21 are rejected under 35 U.S.C. 102(e) as being anticipated by Cook (US 6,631,360).
>
> Claim 1, Cook teaches:
>
> A computer-implemented method of creating customer promotion response models for use in customer relationship marketing, comprising.
>
> (a) generating in a computer an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives (see col 9, lines 35-45; col 12, lines 5-30; "data source, such as buy or no buy data") that are base variables and conditions that are predicates, aggregates or other function where the primitives and conditions determine how the Analytical Variables are derived from operational data to produce the input data set, (col 12, lines 5-30; categories of said data source), and wherein the Analytic Variables are subdivided

into independent variables and their related dependent variables (see col 12, lines 17-22);

(b) splitting in the computer the input data set into a test sample and a validation sample (see col 10, line 55 -col 11, line 20);

(c) identifying in the computer the independent variables and their related dependent variables using the test sample (see col 12, lines 5-45);

(d) identifying in the computer a Transformation Type for each of the identified independent variables and their related dependent variables (see col 11, lines 20-65 "estimated density function");

(e) estimating in the computer a Coefficient for each of the identified independent variables and their related dependent variables (see col 14, lines 55-65 "each element in a decision array there is a gain or loss");

(f) generating in the computer a Model Equation for each of the identified independent variables and their related dependent variables using the identified Transformation Type and estimated Coefficient (see col 13, lines 5-45 "Gaussian Density function");

(g) validating in the computer the generated Model Equation by applying it to the validation sample (see col 11, lines 5-20 "calibration"; and

(h) scoring in the computer customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing (see col 11, lines 50-67).

The Office Action also states the following:

Response to Arguments

4. Applicant's arguments filed 12/04/2008 have been fully considered but they are not persuasive. The Applicant argues that Cook does not teach "generating an input data set for the response model, wherein the input data set is generated using an Analytical Data Set template containing one or more Analytical Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions that describe how the Analytical Variables are derived from operational data to produce the input data set" and wherein the Analytic Variable are subdivided into independent variables and their related dependent variables". The Applicant further argues that Cook does not teach how this data is created, other than by profiling or collecting, that Cook does not teach "Analytical Data Set templates to generate data and that Cook does teach that the Analytical Data Set templates contain analytical variables and that Cook does not generates its input data set from operation data using primitives and conditions of Analytical Variables contained within Analytical Data Set Templates. The Examiner answers that Cook teaches selecting a test sample or training sample and a validating sample (i.e. unknown sample data) from a data source (see col 8, lines 20-25, col 15, lines 1-12) and where said training and validating data contains analytical variables which contains independent and dependent variables. Applicant's specification mentions in page 6 lines 15-32 that "Analytical variables are comprised of primitives and conditions that describe how the Analytical Variable are derived from the

operational data. Primitives are base variables, while conditions are predicates, aggregates or other functions." The Applicant's specification page 6 gives an example, where it recites "for example "Sum of sales" in "Merchandise Department" during "Last 6 months" may identify hundreds of variables. However, the system could create an Analytical Variable by summing a "Sales" base variable (i.e. primitive) associated with multiple primitives (e.g. Department and Transaction Date variables) and conditions (e.g. Department = "Merchandise" and Transaction date > "February 1, 2001").

Thereafter, the user creates an Analytical Data Set Template containing the desired Analytical Variables required for a specific analysis task". Therefore, according to the Applicant's specification, said limitation of "Analytical Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions that describe how the Analytical Variables are derived from operational data" simply means, according to Applicant's specification, selecting the Analytical variables from base variables by applying some type of condition selection to said base variables. Applicant's specification only recites "that conditions are predicates, aggregates or functions" and nothing else. Cook teaches selecting a base variable category (i.e. buyer/non-buyer) and applying some type of selection function to said data, which for example, is "n selected individuals' related data is removed from the training data structure" in order to create Analytical variables to be used in a density function for each category based on the training data structure with the selected individual's data removed" (see col 3, lines 30-40). Cook teaches applying conditions to primitive data (i.e. categories) in order to determine which analytical variables to use in order to predict if buyers/non buyers, therefore, creating an Analytical Set template. Furthermore, Cook teaches that that data source may include independent (i.e. profile features such as buy or not buy) and dependent variables (i.e. category into which a profile individual falls) (see col 12, lines 10-30). Therefore, contrary to Applicant's argument, Cook teaches Applicant's claimed limitation. The Applicant argues that Cook does not teach "splitting the input data set into a test sample and a validation sample. The Examiner answers that Cook teaches selecting (i.e. generating) a test sample (i.e. training sample) and a validation sample (i.e. unknown sample) from a data source (see col 15, lines 1-15; col 8, lines 20-25). Therefore, contrary to Applicant's argument, Cook teaches "splitting" the data source into a training sample and a validation sample.

The Applicant argues that Cook does not teach "identifying independent and their related dependent variables using the test sample". The Examiner answers that Cook teaches identifying independent and dependent variables from a test or training sample (see col 12, lines 15-25). Therefore, contrary to Applicant's argument, Cook teaches Applicant's claimed invention.

The Applicant argues that Cook does not teach "identifying a transformation type, for each of the identified independent and their related dependent variables, where the transformation type provide the strongest association between the identified related independent variable and the dependent variables and generating a Model Equation for each of the identified independent and their related dependent variables using the identified Transformation type and

estimated coefficient". The Examiner answers that Cook teaches probability density functions that result in normal or quadratic decision surfaces (see col 10, lines 1-10), where said density function is used to create a decision array (see col 3, lines 45-55) and where each element of the decision array there is a gain or loss (see col 14, lines 55-65) which shows an association between the identified related independent variables (i.e. individual profile features see col 10, lines 55-65) and the dependent variables (i.e. category into which a profile individual falls) (see col 12, lines 10-30). Therefore, contrary to Applicant's argument, Cook teaches Applicant's claimed limitation.

The Applicant argues that Cook does not teach "estimating a coefficient for the identified related independent and dependent variables". The Examiner answers that Cook figures 12 and 13 teach estimating coefficients (i.e. density value) for each independent and dependent variable of said graph. Therefore, contrary to Applicant's argument, Cook teaches Applicant's claimed limitation.

The Applicant argues that Cook does not teach a transformation type because Cook's density function relates to the "distribution" of independent variables among categories, whereas Applicant's model equations relates to why (mathematical) an associated independent variable is associated with a particular dependent variable. The Examiner answers that the Applicant is arguing about limitation not stated in the claims when he mentions that Applicant's claims recite "why" a variable is associated with another. However, Cook teaches said association between independent and dependent variables in col 12, lines 10-25). Therefore, contrary to Applicant's argument, Cook teaches Applicant's claimed invention.

The Applicant argues that Cook does not teach "validating the generated Model Equation by applying it to validation sample". The Examiner answers that Cook teaches performing a calibration process to determine the accuracy of a forecast (see col 11, lines 5-20). Therefore, contrary to Applicant's argument, Cook teaches Applicant's claimed invention.

The Applicant argues that Cook does not teach "scoring customers retrieved from a database using a Model Equation". The Examiner answers that Cook figures 12 and 13 teach determining the relative density value (i.e. score) for each individual category, feature and category. Therefore, contrary to Applicant's argument, Cook teaches Applicant's claimed limitation.

Appellants' attorney respectfully disagrees with this analysis for the reasons set forth below.

i. <u>The cited portions of Cook do not teach or suggest "generating an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions, wherein the</u>

primitives and conditions determine how the Analytic Variables are derived from operational data to produce the input data set, and wherein the Analytic Variables are subdivided into independent variables and their related dependent variables."

The portions of Cook cited by the Office Action as teaching the above limitations found in Appellants' claims are set forth below:

Cook: col. 8, lines 20-25

The collected user profile data is stored in databases on the Web servers of the businesses collecting the data or databases of other businesses hired to collect the user profile data. The purpose of the collection is to enhance the ability to help customers and to help encourage users to purchase the goods and/or services of the business collecting the user profile data.

Cook: col. 9, lines 35-45

Initially, as shown by block 401, a training sample is set up. As will be better understood from the following description of an example of a training sample setup process, setting up a training sample involves defining and naming categories and identifying a source of data for each defined category, i.e., a source of data that contains profile feature information (i.e., data) regarding individuals that fall in the defined categories. The data source must associate the profile data with the categories. The training sample setup 401 also involves assembling the data into a predetermined data structure.

Cook: col. 12, lines 5-30 (actually, lines 5-28)

FIG. 5 illustrates a training sample setup process 401 formed in accordance with the present invention. Initially, categories are selected 501. As noted above, selecting categories involves defining the categories and naming them, i.e., buyer/non-buyer, responder/non-responder, responder/non-responder/unsubscribe, sick/healthy, friend/foe, etc. Category names are normally entered into a computer system by a user via a graphical user interface (GUI), also called a dialog window. Next, for a selected category a data source is identified. The data source may be as simple as a manually preformatted file. Alternatively, and more likely, the data source is a source of data developed by profiling Internet customers. The data source must include independent variables, i.e., individual profile features and an associated dependent variable, i.e., the category into which a profiled individual falls. The data may be collected, for example, by advertising a product to a selected group of potential purchasers whose profile is known to the advertiser. The buy or no-buy results, combined with the potential purchasers' profile features, creates the data source for the selected category, i.e., buy or no-buy. Next, a test 505 is made to determine if any more categories have been entered by the user. If so, the next category is selected and a data source is identified, which may be the same data source.

> Cook: col. 15, lines 1-12 (actually, col. 14, line 66 – col. 12, line 12)
>
> FIG. 9 illustrates a training data setup process 421 formed in accordance with the invention. Initially, a data source is identified 901. This may be as simple as identifying a manually preformatted file, or this may involve an arbitrary complex database extraction. Typically, the extraction process is the same as for the training sample setup, except that category data is not required, only individual profile feature data. Next, as with training sample data, the extracted data is assembled into a computer file in a specific format and stored in a workplace (temporary memory) 903. Thereafter, the extracted assembled data is used to create an unknown sample data structure and the structure is stored. An example of an unknown data structure formed in accordance with this invention is illustrated in FIG. 14.

The above portions of Cook do not teach or suggest "generating an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions, wherein the primitives and conditions determine how the Analytic Variables are derived from operational data to produce the input data set, and wherein the Analytic Variables are subdivided into independent variables and their related dependent variables."

Nonetheless, in the "Response to Arguments," the Office Action analyzes Applicant's specification as it relates to "Analytical Variables," and notes that Analytical Variables are comprised of primitives and conditions that describe how the Analytical Variable are derived from the operational data. The Office Action then asserts that Cook teaches this limitation because it selects a base variable category (i.e. buyer/non-buyer) and applies some type of selection function to the data, or applies conditions to primitive data (i.e. categories), or includes data sources with independent and dependent variables. From this, the Office Action asserts that Cook teaches Applicant's claimed limitation.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

Specifically, nowhere does Cook teach or suggest Analytic Data Set Templates. In this regard, Cook does not use Analytic Data Set Templates to generate data, Cook does not teach that Analytic Data Set Templates contain Analytic Variables, and Cook does not generate its input data sets from operational data using primitives and conditions of Analytic Variables contained within Analytic Data Set Templates.

Instead, the above portions of Cook merely describe data that contains profile feature information (e.g., independent variables) regarding individuals that fall in the defined categories (e.g., dependent variables). However, Cook does not describe how this data is created, other than by "profiling" or "collecting," and merely states that "[t]he data source must include independent variables, i.e., individual profile features and an associated dependent variable, i.e., the category into which a profiled individual falls."

ii. <u>The cited portions of Cook do not teach or suggest "splitting the input data set into a test sample and a validation sample."</u>

The portions of Cook cited by the Office Action as teaching the above limitations found in Appellants' claims are set forth below:

> <u>Cook: col. 8, lines 20-25</u>
> The collected user profile data is stored in databases on the Web servers of the businesses collecting the data or databases of other businesses hired to collect the user profile data. The purpose of the collection is to enhance the ability to help customers and to help encourage users to purchase the goods and/or services of the business collecting the user profile data.
>
> <u>Cook: col. 10, line 55 – col. 11, line 20</u>
> Returning to FIG. 4A, after the first inference engine is selected, a set of profile features are selected 409. This step is included so that independent variables (individual profile features) that are not different among categories can be eliminated. Preferably individual profile features are sorted based on standard statistics after controlling for multicolinearities. In addition to eliminating independent variables for which there is insufficient data for estimation, less significant individual profile features can also be eliminated if desired. The end result is one or more sets of profile features. At step 409 one set is selected.
>
> After the inference engine and set of profile features have been selected, a training process is conducted 411. An example of a training process formed in accordance with the invention is illustrated in FIG. 6 and described below. In general, during the training process, various probability density functions are estimated for the selected engine and a data structure containing unbiased density values is created.
>
> After the training process 411 is completed, a calibration process 413 is performed. An example of a calibration process formed in accordance with the invention is illustrated in FIG. 8 and described below. The calibration process creates a decision array in which are stored the results of classifying the individuals whose individual profile features were contained in the training sample. As will be better understood from the following description, the decision array compares an individual's true category to the category predicted by the

selected inference engine. The decision array in combination with the estimated density function and density value data structure contain all the algorithms and parameters necessary for implementation of the selected engine.

<u>Cook: col. 15, lines 1-12 (actually, col. 14, line 66 – col. 12, line 12)</u>

FIG. 9 illustrates a training data setup process 421 formed in accordance with the invention. Initially, a data source is identified 901. This may be as simple as identifying a manually preformatted file, or this may involve an arbitrary complex database extraction. Typically, the extraction process is the same as for the training sample setup, except that category data is not required, only individual profile feature data. Next, as with training sample data, the extracted data is assembled into a computer file in a specific format and stored in a workplace (temporary memory) 903. Thereafter, the extracted assembled data is used to create an unknown sample data structure and the structure is stored. An example of an unknown data structure formed in accordance with this invention is illustrated in FIG. 14.

Consider also the following pertinent portions of Cook:

<u>Cook: col. 11, line 20 - col. 12, line 4</u>

After the calibration process has been completed for the selected engine, a test 415 is made to determine if any more sets of features need to be processed for the selected inference engine. If so, the next set of features is selected and the training and calibration processes 411 and 413 are repeated. If no more sets of features need to be processed, the set of features for the selected engine that best meets the desired objective are selected 417. Then a test is made to determine if any additional inference engines are to be selected. If so, as shown by decision block 415, the foregoing process is repeated, i.e., another inference engine is selected 409, a set of features is selected 409, and the training process is performed 411, followed by the calibration process 413. The foregoing sequence is repeated until no more inference engines remain to be selected.

Preferably, as each inference engine is processed, the resulting decision array is analyzed during the calibration step to determine if the current inference engine is better than previously processed inference engines. If so, the array and the estimated density function for the current inference engine replace previously stored decision array and estimated density function data. If the current inference engine is not better, the previously stored decision array and the estimated density function data is retained and this data for the current inference engine is discarded. Thus, after the training and calibration processes are complete for all inference engines, the decision array and estimated density function for the best inference engine are stored.

After the best inference engine has been selected in the foregoing manner, a sample comprising individual observations for which category membership is unknown is identified. The unknown sample contains the same independent variables (individual profile features) as did the training sample and is set up 421 in generally the same manner as the training sample was set up 401. An example

> of an unknown sample setup process formed in accordance with this invention is illustrated in FIG. 9 and described below. Thereafter, each individual in the unknown sample is assigned to a category using the previously developed and stored estimated density function associated with the selected best inference engine. Each such assignment is called a prediction. The predictions are tallied and the tally adjusted for error rates determined by the decision array created during the calibration process described above. The result is a forecast. See block 423. Next, a test 425 is made to determine if another unknown sample is to be analyzed. If so, the foregoing steps are repeated. After all unknown samples have been examined, the user can determine 427 if the objective needs to be reset. If the objective needs to be reset, the objective is reset and the entire process is repeated. If not, the process ends.

The above portions of Cook do not teach or suggest "splitting the input data set into a test sample and a validation sample." Nonetheless, in the "Response to Arguments," the Office Action asserts that the above portions of Cook teach selecting (i.e. generating) a test sample (i.e. training sample) and a validation sample (i.e. unknown sample) from a data source, and that this comprises "splitting" the data source into a training sample and a validation sample.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

For example, there is no "validation sample" in the above portions of Cook. Instead, the unknown sample from FIG. 9 is used to create predictions and forecasts. Note too that FIG. 9 refers to a "training data setup process," not a validation process. Moreover, nowhere do the above portions of Cook describe splitting a data set. Instead, the above portions of Cook describe a training sample being developed independently, for example, of any validation sample.

iii. <u>The cited portions of Cook do not teach or suggest "identifying independent and their related dependent variables using the test sample."</u>

The portions of Cook cited by the Office Action as teaching the above limitations found in Appellants' claims are set forth below:

> <u>Cook: col. 12, lines 5-45</u>
>
> FIG. 5 illustrates a training sample setup process 401 formed in accordance with the present invention. Initially, categories are selected 501. As noted above, selecting categories involves defining the categories and naming them, i.e., buyer/non-buyer, responder/non-responder, responder/non-responder/unsubscribe, sick/healthy, friend/foe, etc. Category names are normally entered into a computer system by a user via a graphical user interface (GUI), also

> called a dialog window. Next, for a selected category a data source is identified. The data source may be as simple as a manually preformatted file. Alternatively, and more likely, the data source is a source of data developed by profiling Internet customers. The data source must include independent variables, i.e., individual profile features and an associated dependent variable, i.e., the category into which a profiled individual falls. The data may be collected, for example, by advertising a product to a selected group of potential purchasers whose profile is known to the advertiser. The buy or no-buy results, combined with the potential purchasers' profile features, creates the data source for the selected category, i.e., buy or no-buy. Next, a test 505 is made to determine if any more categories have been entered by the user. If so, the next category is selected and a data source is identified, which may be the same data source.
>
> After the data sources have been identified, a training sample set is established 507. This involves downloading data from the data source(s) and assembling the data into a computer file, or set of computer files, in a specific format, and storing the files in a workspace, i.e., in temporary memory. After establishing a training set in this matter, the downloaded data is converted into a training data structure having a predetermined configuration and the training data structure is stored in memory 509. A suitable training data structure is illustrated in FIG. 11. The profile features 1111a, 1111b, 1111c . . . 1111n of each individual 1113 in each category 1115 are included in the training data structure. As illustrated in FIG. 4A and discussed above, after the training data structure has been created, selected features of individuals may be eliminated. See step 409, FIG. 4A.

The above portions of Cook do not teach or suggest "identifying independent and their related dependent variables using the test sample." Nonetheless, in the "Response to Arguments," the Office Action asserts that the above portions of Cook teach identifying independent and dependent variables from a test or training sample.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

The above portions of Cook merely describe setting up a training sample by defining the categories, identifying a data source for a selected category (where the data source "must include" both independent variables, i.e., individual profile features and an associated dependent variable, i.e., the category into which a profiled individual falls), and then downloading data from the data source to establish a training set.

However, there is no identification of independent and their related dependent variables being performed in the above portions of Cook. Instead, Cook merely states that the data source "must include" independent variables, i.e., individual profile features and associated dependent

variables, i.e., the category into which a profiled individual falls, without stating how such data is created, other than by "profiling" or "collecting."

iv. <u>The cited portions of Cook do not teach or suggest "identifying a Transformation Type for each of the identified independent and their related dependent variables."</u>

The portions of Cook cited by the Office Action as teaching the above limitations found in Appellants' claims are set forth below:

> <u>Cook: col. 3, lines 45-55 (actually, lines 32-55)</u>
>
> In accordance with further aspects of this invention, the training data structure is analyzed using a leaving-n-out approach. First a category is selected. Next, n selected individuals' related data is removed from the training data structure. Then, a density function is estimated and a density value is calculated for each category based on the training data structure with the selected individual's data removed. The selected individual's data is reinserted into the training data structure and the foregoing sequence is repeated for another individual. After a density function has been estimated and a density value calculated for each category for each individual with the individual's data removed, the entire sequence is repeated for the next category. Processing continues until all categories have been analyzed. The value of n can vary from 1 up to half of the individuals comprising the training data structure.
>
> In accordance with still further aspects of this invention, the calculated density values are used to create a density value data structure. Then, for each category and each individual, a decision rule is applied to the density value data structure for the individual. The results are used to create a decision array. After completion, the decision array is displayed so that a user can determine if the objective has been met.
>
> <u>Cook: col. 10, lines 1-10 (actually, col. 9, line 56 – col. 10, line 10)</u>
>
> After the objective has been set, a first inference engine is selected at 407. As will be better understood from the following description, the invention is architected with Bayes Rule as a framework. This allows any "inference engine" to be formalized in the same context. Bayes Rule effectively says that for a particular individual observation, that observation should be assigned to the category to which the observation has the maximum probability of belonging. The values of the independent variables, i.e., the individual profile features, are used to calculate these probabilities using a variety of inference engines. The inference engines are, in effect, algorithms that make the assumption that independent variables for a given category are distributed according to some probability density function. The most accurate inference engine will typically be the one for which the data are most closely modeled by the assumed probability density

function. The presently preferred probability density functions are (a) normal with equal variances among categories that results in a linear decision surface, (b) normal with unequal variances among categories that results in a quadratic decision surface, and (c) Parzen that results in a polynomial decision surface.

<u>Cook: col. 10, lines 55-65 (actually, col. 9, line 56 – col. 10, line 10)</u>

Returning to FIG. 4A, after the first inference engine is selected, a set of profile features are selected 409. This step is included so that independent variables (individual profile features) that are not different among categories can be eliminated. Preferably individual profile features are sorted based on standard statistics after controlling for multicolinearities. In addition to eliminating independent variables for which there is insufficient data for estimation, less significant individual profile features can also be eliminated if desired. The end result is one or more sets of profile features. At step 409 one set is selected.

<u>Cook: col. 11, lines 20-65 (actually, col. 11, line 20 - col. 12, line 4)</u>

After the calibration process has been completed for the selected engine, a test 415 is made to determine if any more sets of features need to be processed for the selected inference engine. If so, the next set of features is selected and the training and calibration processes 411 and 413 are repeated. If no more sets of features need to be processed, the set of features for the selected engine that best meets the desired objective are selected 417. Then a test is made to determine if any additional inference engines are to be selected. If so, as shown by decision block 415, the foregoing process is repeated, i.e., another inference engine is selected 409, a set of features is selected 409, and the training process is performed 411, followed by the calibration process 413. The foregoing sequence is repeated until no more inference engines remain to be selected.

Preferably, as each inference engine is processed, the resulting decision array is analyzed during the calibration step to determine if the current inference engine is better than previously processed inference engines. If so, the array and the estimated density function for the current inference engine replace previously stored decision array and estimated density function data. If the current inference engine is not better, the previously stored decision array and the estimated density function data is retained and this data for the current inference engine is discarded. Thus, after the training and calibration processes are complete for all inference engines, the decision array and estimated density function for the best inference engine are stored.

After the best inference engine has been selected in the foregoing manner, a sample comprising individual observations for which category membership is unknown is identified. The unknown sample contains the same independent variables (individual profile features) as did the training sample and is set up 421 in generally the same manner as the training sample was set up 401. An example of an unknown sample setup process formed in accordance with this invention is illustrated in FIG. 9 and described below. Thereafter, each individual in the unknown sample is assigned to a category using the previously developed and stored estimated density function associated with the selected best inference

engine. Each such assignment is called a prediction. The predictions are tallied and the tally adjusted for error rates determined by the decision array created during the calibration process described above. The result is a forecast. See block 423. Next, a test 425 is made to determine if another unknown sample is to be analyzed. If so, the foregoing steps are repeated. After all unknown samples have been examined, the user can determine 427 if the objective needs to be reset. If the objective needs to be reset, the objective is reset and the entire process is repeated. If not, the process ends.

<u>Cook: col. 12, lines 10-30 (actually, col. 12, lines 5–43)</u>

FIG. 5 illustrates a training sample setup process 401 formed in accordance with the present invention. Initially, categories are selected 501. As noted above, selecting categories involves defining the categories and naming them, i.e., buyer/non-buyer, responder/non-responder, responder/non-responder/unsubscribe, sick/healthy, friend/foe, etc. Category names are normally entered into a computer system by a user via a graphical user interface (GUI), also called a dialog window. Next, for a selected category a data source is identified. The data source may be as simple as a manually preformatted file. Alternatively, and more likely, the data source is a source of data developed by profiling Internet customers. The data source must include independent variables, i.e., individual profile features and an associated dependent variable, i.e., the category into which a profiled individual falls. The data may be collected, for example, by advertising a product to a selected group of potential purchasers whose profile is known to the advertiser. The buy or no-buy results, combined with the potential purchasers' profile features, creates the data source for the selected category, i.e., buy or no-buy. Next, a test 505 is made to determine if any more categories have been entered by the user. If so, the next category is selected and a data source is identified, which may be the same data source.

After the data sources have been identified, a training sample set is established 507. This involves downloading data from the data source(s) and assembling the data into a computer file, or set of computer files, in a specific format, and storing the files in a workspace, i.e., in temporary memory. After establishing a training set in this matter, the downloaded data is converted into a training data structure having a predetermined configuration and the training data structure is stored in memory 509. A suitable training data structure is illustrated in FIG. 11. The profile features 1111a, 1111b, 1111c . . . 1111n of each individual 1113 in each category 1115 are included in the training data structure. As illustrated in FIG. 4A and discussed above, after the training data structure has been created, selected features of individuals may be eliminated. See step 409, FIG. 4A.

<u>Cook: col. 14, lines 55-65</u>

As will be readily appreciated by those skilled in the art to which this invention pertains, in the real world, each category will have associated benefits and costs. These benefits may be tangible, as in the case of dollars, or intangible, as in the case of goodwill. Thus, for each element of the decision array there is a

> gain or loss that can be assigned to each individual within that element. The net gain or loss for each element of the decision array is the individual gain or loss multiplied by the number of individuals assigned to that element. The objective function is thus the sum of these net gains or losses.

The above portions of Cook do not teach or suggest "identifying a Transformation Type for each of the identified independent and their related dependent variables." Nonetheless, in the "Response to Arguments," the Office Action asserts that the above portions of Cook teach probability density functions that result in normal or quadratic decision surfaces, where the density function is used to create a decision array and where each element of the decision array is a gain or loss that shows an association between the identified related independent variables (i.e. individual profile features) and the dependent variables (i.e. category into which a profile individual falls).

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

As noted in Appellants' specification, a Response Modeling service identifies a Transformation Type for the identified related independent and dependent variables, i.e., the predictive variables. The Transformation Type is a mathematical operation that provides the strongest association between the identified related independent variable and the dependent variables, namely "why" (mathematically) an independent variable is associated with a particular dependent variable.

Instead, the above portions of Cook merely describe the inference engines as algorithms that make the assumption that independent variables for a given category are distributed according to some probability density function.

However, the probability density function of Cook is not a Transformation Type. A Transformation Type is defined as a mathematical operation that provides the strongest "association" between the identified independent variables and their related dependent variables. Cook's probability density function, on the other hand, relates to the "distribution" of independent variables among different categories, not the "association" between one or more of the independent variables and their specific related category.

v. <u>The cited portions of Cook do not teach or suggest "estimating a Coefficient for each of the identified independent and their related dependent variables."</u>

The portions of Cook cited by the Office Action as teaching the above limitations found in Appellants' claims are set forth below:

<u>Cook: FIGS. 12-14</u>

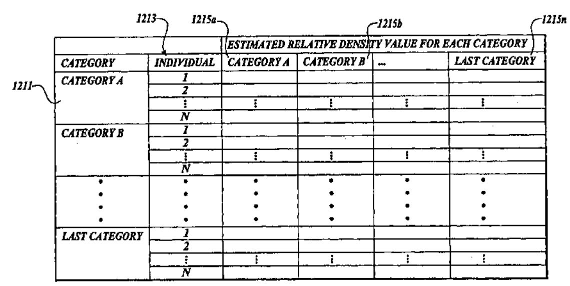

Fig. 12

U.S. Patent Oct. 7, 2003 Sheet 13 of 14 US 6,631,360 B1

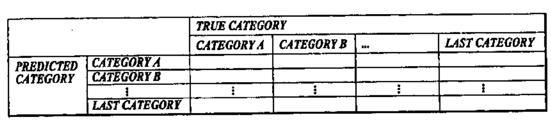

Fig. 13

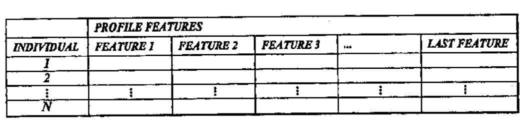

Fig. 14

U.S. Patent Oct. 7, 2003 Sheet 14 of 14 US 6,631,360 B1

<u>Cook: col. 14, lines 55-65</u>

As will be readily appreciated by those skilled in the art to which this invention pertains, in the real world, each category will have associated benefits and costs. These benefits may be tangible, as in the case of dollars, or intangible, as in the case of goodwill. Thus, for each element of the decision array there is a gain or loss that can be assigned to each individual within that element. The net gain or loss for each element of the decision array is the individual gain or loss

> multiplied by the number of individuals assigned to that element. The objective function is thus the sum of these net gains or losses.

The above portions of Cook do not teach or suggest "estimating a Coefficient for each of the identified independent and their related dependent variables." Nonetheless, in the "Response to Arguments," the Office Action asserts that the above portions of Cook teach estimating coefficients (i.e. density value) for each independent and dependent variable of the graphs of FIGS. 12 and 13.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

As noted in Appellants' specification, a Response Modeling service identifies a Transformation Type for the identified related independent and dependent variables, i.e., the predictive variables. The Transformation Type is a mathematical operation that provides the strongest association between the identified related independent variable and the dependent variables.

After identifying a Transformation Type, the Response Modeling service estimates a Coefficient, or weight, for each of the identified related independent and dependent variables found to be significant in predicting the likelihood of response. The Coefficient is a relative measure of the contribution of a variable to the likelihood of response. However, the size of the Coefficient does not indicate the relative importance of the variable in predicting the likelihood of response, since it is itself dependent on the magnitude of the variable. The sign of the Coefficient indicates whether the independent variable is positively or negatively correlated with the dependent variable.

Instead, the above portions of Cook merely describe estimated relative density values for each category, namely the frequency of occurrence of independent variables in the categories. However, the estimated relative density values of Cook are not Coefficients, which are defined as a relative measure (e.g., a weight) of the identified independent and their related dependent variables' contributions to a likelihood of response.

vi. <u>The cited portions of Cook do not teach or suggest "generating a Model Equation for each of the identified independent and their related dependent variables using the identified Transformation Type and estimated Coefficient."</u>

The portions of Cook cited by the Office Action as teaching the above limitations found in Appellants' claims are set forth below:

> <u>Cook: col. 12, lines 10-25 (actually, col. 12, lines 5-43)</u>
>
> FIG. 5 illustrates a training sample setup process 401 formed in accordance with the present invention. Initially, categories are selected 501. As noted above, selecting categories involves defining the categories and naming them, i.e., buyer/non-buyer, responder/non-responder, responder/non-responder/unsubscribe, sick/healthy, friend/foe, etc. Category names are normally entered into a computer system by a user via a graphical user interface (GUI), also called a dialog window. Next, for a selected category a data source is identified. The data source may be as simple as a manually preformatted file. Alternatively, and more likely, the data source is a source of data developed by profiling Internet customers. The data source must include independent variables, i.e., individual profile features and an associated dependent variable, i.e., the category into which a profiled individual falls. The data may be collected, for example, by advertising a product to a selected group of potential purchasers whose profile is known to the advertiser. The buy or no-buy results, combined with the potential purchasers' profile features, creates the data source for the selected category, i.e., buy or no-buy. Next, a test 505 is made to determine if any more categories have been entered by the user. If so, the next category is selected and a data source is identified, which may be the same data source.
>
> After the data sources have been identified, a training sample set is established 507. This involves downloading data from the data source(s) and assembling the data into a computer file, or set of computer files, in a specific format, and storing the files in a workspace, i.e., in temporary memory. After establishing a training set in this matter, the downloaded data is converted into a training data structure having a predetermined configuration and the training data structure is stored in memory 509. A suitable training data structure is illustrated in FIG. 11. The profile features 1111a, 1111b, 1111c . . . 1111n of each individual 1113 in each category 1115 are included in the training data structure. As illustrated in FIG. 4A and discussed above, after the training data structure has been created, selected features of individuals may be eliminated. See step 409, FIG. 4A.
>
> <u>Cook: col. 13, lines 5-45</u>
>
> As noted above, FIG. 7 illustrates a density function and density value calculating process 605 suitable for use in the process illustrated in FIG. 6 formed in accordance with the invention. First, a category is selected 701 by, for example,

> setting a pointer to the memory location of the data associated with the selection-- in this case, the first category. Then, the density function for the category is estimated 703. More specifically, the parameters for the density function for the selected category are estimated from the training data structure (FIG. 11). For example, in the case where a Gaussian density function is used, the mean for each selected feature (FIG. 11) and the variance-covariance matrix for these features are estimated within each category (FIG. 11). These estimates become the parameter values in the estimated Gaussian density function for each category. In this estimated Gaussian density function, there exists a variable for each selected feature. The thusly created estimated density function is stored 705. Then, the estimated density function for the selected category is used to calculate an estimated relative density value for the selected individual in the selected category 707. More specifically, using the foregoing example, the values of the selected features are substituted for the variables in the estimated Gaussian density function and a scalar is obtained. The result is used to create a density value data structure, which is stored 709. Then a test 711 is made to determine if any more categories exist. If more categories exist, the process is repeated. As will be recalled, the process illustrated in FIG. 7 occurs after the training data structure has been updated by removing a selected individual from a selected category. Thus, the density value data structure is for a selected individual in a selected category with the n individuals' data removed. An example of a density value data structure is shown in FIG. 12. For each category 1211 an estimate of the likelihood that each individual 1213 will fall in each category 1215a, 1215b . . . 1215n is included in the data structure.

The above portions of Cook do not teach or suggest "generating a Model Equation for each of the identified independent and their related dependent variables using the identified Transformation Type and estimated Coefficient." Nonetheless, in the "Response to Arguments," the Office Action asserts that the Applicant is arguing about limitation not stated in the claims when he mentions that Applicant's claims recite "why" a variable is associated with another, and that the above portions of Cook teach an association between independent and dependent variables.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

For example, the above portions of Cook merely describe inference engines as algorithms that make the assumption that independent variables for a given category are distributed according to some probability density function, and that the best inference engine is determined, using the training sample, based on an estimated Gaussian density function. As described in Cook, the estimated Gaussian density function estimates the proportions of selected

subpopulations in a larger population, e.g., the frequency of occurrence of the independent variable in a category.

In addition, the above portions of Cook merely describe that a category is selected, the parameters for the density function for the selected category are estimated from the training data (in the case where a Gaussian density function is used, the mean for each selected feature and the variance-covariance matrix for these features are estimated within each category), and then the estimated density function for the selected category is used to calculate an estimated relative density value for a selected individual in the selected category (the values of the selected features are substituted for the variables in the estimated Gaussian density function and a scalar is obtained).

The density function of Cook describes the "density" of a variable at a point, e.g., the frequency of occurrence of a variable at a point. Cook's density function therefore relates to the "distribution" of independent variables among categories (dependent variables). However, the distribution of independent variables among dependent variables is not a mathematical representation of the association of independent variables and their related dependent variables.

As noted in Appellants' specification, a Response Modeling service identifies a Transformation Type for the identified related independent and dependent variables, i.e., the predictive variables. The Transformation Type is a mathematical operation that provides the strongest association between the identified related independent variable and the dependent variables.

After identifying a Transformation Type, the Response Modeling service estimates a Coefficient, or weight, for each of the identified related independent and dependent variables found to be significant in predicting the likelihood of response. The Coefficient is a relative measure of the contribution of a variable to the likelihood of response. However, the size of the Coefficient does not indicate the relative importance of the variable in predicting the likelihood of response, since it is itself dependent on the magnitude of the variable. The sign of the Coefficient indicates whether the independent variable is positively or negatively correlated with the dependent variable.

Finally, after estimating a Coefficient, the Response Modeling service generates a Model Equation that is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders

versus non-responders. Specifically, the Model Equation includes an association of the independent variable with the dependent variable that best differentiates responders from non-responders, as well as the Transformation Type and the Coefficients associated with the variables.

Cook's density function, which relates to the "distribution" of independent variables among categories (dependent variables), would not comprise a Transformation Type, which is a mathematical operation that provides the strongest association between the identified related independent variable and the dependent variables, or a Coefficient, which is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders. Consequently, Cook does not teach or suggest "generating a Model Equation using a Transformation Type and a Coefficient."

vii. The cited portions of Cook do not teach or suggest "validating the generated Model Equation by applying it to the validation sample."

The portions of Cook cited by the Office Action as teaching the above limitations found in Appellants' claims are set forth below:

> Cook: col. 11, lines 5-20
>
> After the training process 411 is completed, a calibration process 413 is performed. An example of a calibration process formed in accordance with the invention is illustrated in FIG. 8 and described below. The calibration process creates a decision array in which are stored the results of classifying the individuals whose individual profile features were contained in the training sample. As will be better understood from the following description, the decision array compares an individual's true category to the category predicted by the selected inference engine. The decision array in combination with the estimated density function and density value data structure contain all the algorithms and parameters necessary for implementation of the selected engine.

The above portions of Cook do not teach or suggest "validating the generated Model Equation by applying it to the validation sample." Nonetheless, in the "Response to Arguments," the Office Action asserts that the above portions of Cook teach performing a calibration process to determine the accuracy of a forecast.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

Calibration in Cook refers to a process that creates a decision array from the results of the training sample. Calibration generally is considered a process for establishing a relationship between a measuring device and the units of measure, in order to quantify an uncertainty estimate.

Validation in Appellants' invention applies a Model Equation to a validation sample, which is created by splitting the input data set, after the Model Equation has been generated, using of the Transformation Type and Coefficient identified from the independent and dependent variables of a test sample. As described in Appellants' specification, validation of the Model Equation compares a predicted likelihood of response with an actual response.

viii. <u>The cited portions of Cook do not teach or suggest "scoring customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing."</u>

The portions of Cook cited by the Office Action as teaching the above limitations found in Appellants' claims are set forth below:

<u>Cook: FIGS. 12-14</u>

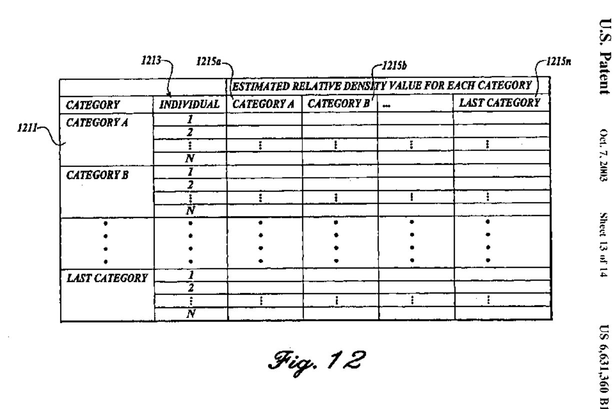

Fig. 12

		TRUE CATEGORY			
		CATEGORY A	CATEGORY B	...	LAST CATEGORY
PREDICTED CATEGORY	CATEGORY A				
	CATEGORY B				
	⋮	⋮	⋮	⋮	⋮
	LAST CATEGORY				

Fig. 13

	PROFILE FEATURES				
INDIVIDUAL	FEATURE 1	FEATURE 2	FEATURE 3	...	LAST FEATURE
1					
2					
⋮	⋮	⋮	⋮	⋮	⋮
N					

Fig. 14

<u>Cook: col. 11, lines 50-67 (actually, col. 11, line 49 – col. 12, line 4)</u>

> After the best inference engine has been selected in the foregoing manner, a sample comprising individual observations for which category membership is unknown is identified. The unknown sample contains the same independent variables (individual profile features) as did the training sample and is set up 421 in generally the same manner as the training sample was set up 401. An example of an unknown sample setup process formed in accordance with this invention is illustrated in FIG. 9 and described below. Thereafter, each individual in the unknown sample is assigned to a category using the previously developed and stored estimated density function associated with the selected best inference engine. Each such assignment is called a prediction. The predictions are tallied and the tally adjusted for error rates determined by the decision array created during the calibration process described above. The result is a forecast. See block 423. Next, a test 425 is made to determine if another unknown sample is to be analyzed. If so, the foregoing steps are repeated. After all unknown samples have been examined, the user can determine 427 if the objective needs to be reset. If the objective needs to be reset, the objective is reset and the entire process is repeated. If not, the process ends.

The above portions of Cook do not teach or suggest "scoring customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing." Nonetheless, in the "Response to Arguments," the Office Action asserts that the above portions of Cook teach determining the relative density value (i.e. score) for each individual category, feature and category.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

The above portions of Cook merely describe using the estimated density function for a selected category to calculate an estimated relative density value for a selected individual in the selected category. As noted previously, the estimated density function relates to the

"distribution" of independent variables among categories (dependent variables). As defined in Appellants' specification, a Model Equation is a mathematical representation of the association of independent variables and their related dependent variables, namely "why" (mathematically) an independent variable is associated with a particular dependent variable. Consequently, Cook does not describe the same Model Equation as recited in Appellants' claims.

ix. Summary: Appellants' claimed invention is patentable over Cook.

In light of the above, Appellants' attorney submits that independent claims 1, 8, and 15 are allowable over Cook. Further, dependent claims 2-7, 9-14, and 16-21 are submitted to be allowable over Cook in the same manner, because they are dependent on independent claims 1, 8, and 15, respectively, and thus contain all the limitations of the independent claims. In addition, dependent claims 2-7, 9-14, and 16-21 recite additional novel elements not shown by Cook.

2. Dependent claims 2, 9 and 16

With regard to dependent claims 2, 9 and 16, which recite that "the Transformation Type is a mathematical operation that identifies an association between the identified related independent and dependent variables," the Office Action asserts that these limitations are shown in Cook at col. 12, lines 5-45.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

As noted above, the cited portions of Cook merely describe a training sample setup process, and a probability density function that relates to the distribution of independent variables among categories. However, nowhere do the cited portions of Cook teach or suggest a Transformation Type that identifies an association between the identified related independent and dependent variables.

3. Dependent claims 3, 10 and 17

With regard to dependent claims 3, 10 and 17, which recite that "the Coefficient is a relative measure of the identified related independent and dependent variables' contributions to a likelihood of response," the Office Action asserts that these limitations are shown in Cook at col. 12, lines 5-20 and col. 13, lines 25-45.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

Cook merely describes estimated relative density values for each category, namely the frequency of occurrence of independent variables in the categories. In Appellants' invention, however, a Coefficient, or weight, is estimated for each of the identified related independent and dependent variables found to be significant in predicting the likelihood of response, wherein the Coefficient is a relative measure of the contribution of a variable to the likelihood of response. Thus, the estimated relative density values of Cook are not Coefficients as defined in Appellants' invention.

4. <u>Dependent claims 4, 11 and 18</u>

With regard to dependent claims 4, 11 and 18, which recite that "the Coefficient's sign indicates whether the independent variable is positively or negatively correlated with the dependent variable," the Office Action asserts that these limitations are shown in Cook at col. 14, lines 55-65.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

Cook merely describes estimated relative density values for each category, namely the frequency of occurrence of independent variables in the categories. As noted above, in Appellants' invention, a Coefficient is a weight that is estimated for each of the identified related independent and dependent variables found to be significant in predicting the likelihood of response, wherein the Coefficient is a relative measure of the contribution of a variable to the likelihood of response. Thus, the estimated relative density values of Cook are not Coefficients as defined in Appellants' invention.

5. <u>Dependent claims 5, 12 and 19</u>

With regard to dependent claims 5, 12 and 19, which recite that "the Model Equation is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders," the Office Action asserts that these limitations are shown in Cook at col. 12, lines 5-12.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

Cook's density function, which relates to the "distribution" of independent variables among categories (dependent variables), while Appellants' Model Equation is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders.

6. Dependent claims 6, 13 and 20

With regard to dependent claims 6, 13 and 20, which recite that "the … validating … further comprises … applying the generated Model Equation to the validation sample in order to predict a likelihood of response as compared to an actual response in the validation sample," the Office Action asserts that these limitations are shown in Cook at col. 11, lines 5-20 and col. 13, lines 5-45.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

Cook describes a calibration process that creates a decision array from the results of the training sample, and calibration generally is considered a process for establishing a relationship between a measuring device and the units of measure, in order to quantify an uncertainty estimate. Appellants' invention, on the other hand, applies a Model Equation to a validation sample, which is created by splitting the input data set, after the Model Equation has been generated, using of the Transformation Type and Coefficient identified from the independent and dependent variables of a test sample. As described in Appellants' specification, validation of the Model Equation compares a predicted likelihood of response with an actual response, which is a different from Cook's calibration process.

7. Dependent claims 7, 14 and 21

With regard to dependent claims 7, 14 and 21, which recite that "the … scoring … further comprises … applying the validated Model Equation to the customers retrieved from the database in order to predict responses from the customers in a future promotional campaign," the Office Action asserts that these limitations are shown in Cook at col. 11, lines 50-65 and col. 13, lines 5-45.

Appellants' attorney respectfully submits that the Office Action errs in its analysis.

Cook merely describe using the estimated density function for a selected category to calculate an estimated relative density value for a selected individual in the selected category.

As noted previously, the estimated density function relates to the "distribution" of independent variables among categories (dependent variables). In contrast, Appellants' invention applies a validated Model Equation, which is a mathematical representation of the association of independent variables and their related dependent variables, to customers retrieved from a database in order to predict responses from the customers in a future promotional campaign. Consequently, Cook does not describe the same use of a Model Equation as recited in Appellants' claims.

VIII. CONCLUSION

In light of the above arguments, Appellants' attorney respectfully submits that the cited reference does not anticipate nor render obvious the claimed invention. More specifically, Appellants' claims recite novel physical features which patentably distinguish over the cited reference under 35 U.S.C. §§ 102 and 103.

As a result, a decision by the Board of Patent Appeals and Interferences reversing the Examiner and directing allowance of the pending claims in the subject application is respectfully solicited.

Respectfully submitted,

GATES & COOPER LLP
Attorneys for Appellants

Howard Hughes Center
6701 Center Drive West, Suite 1050
Los Angeles, California 90045
(310) 641-8797

Date: August 11, 2009

By: /George H. Gates/
Name: George H. Gates
Reg. No.: 33,500

GHG/

CLAIMS APPENDIX

1. (PREVIOUSLY PRESENTED) A computer-implemented method of creating a customer promotion response model for use in customer relationship marketing, comprising:

(a) generating, in a computer, an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions, wherein the primitives and conditions determine how the Analytic Variables are derived from operational data to produce the input data set, and wherein the Analytic Variables are subdivided into independent variables and their related dependent variables;

(b) splitting, in the computer, the input data set into a test sample and a validation sample;

(c) identifying, in the computer, the independent variables and their related dependent variables using the test sample;

(d) identifying, in the computer, a Transformation Type for each of the identified independent variables and their related dependent variables;

(e) estimating, in the computer, a Coefficient for each of the identified independent variables and their related dependent variables;

(f) generating, in the computer, a Model Equation for each of the identified independent variables and their related dependent variables using the identified Transformation Type and estimated Coefficient;

(g) validating, in the computer, the generated Model Equation by applying it to the validation sample; and

(h) scoring, in the computer, customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing.

2. (ORIGINAL) The method of claim 1, wherein the Transformation Type is a mathematical operation that identifies an association between the identified related independent and dependent variables.

3. (ORIGINAL) The method of claim 1, wherein the Coefficient is a relative measure of the identified related independent and dependent variables' contributions to a likelihood of response.

4. (ORIGINAL) The method of claim 1, wherein the Coefficient's sign indicates whether the independent variable is positively or negatively correlated with the dependent variable.

5. (PREVIOUSLY PRESENTED) The method of claim 1, wherein the Model Equation is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders.

6. (ORIGINAL) The method of claim 1, wherein the validating step (g) further comprises applying the generated Model Equation to the validation sample in order to predict a likelihood of response as compared to an actual response in the validation sample.

7. (ORIGINAL) The method of claim 1, wherein the scoring step (h) further comprises applying the validated Model Equation to the customers retrieved from the database in order to predict responses from the customers in a future promotional campaign.

8. (PREVIOUSLY PRESENTED) A computer-implemented system for creating a customer promotion response model for use in customer relationship marketing, comprising:

(a) a computer;

(b) a customer relationship marketing system, performed by the computer, for:

(1) generating an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions, wherein the primitives and conditions determine how the Analytic Variables are derived from operational data to produce the input data set, and wherein the Analytic Variables are subdivided into independent variables and their related dependent variables;

(2) splitting the input data set into a test sample and a validation sample;

(3) identifying the independent variables and their related dependent variables using the test sample;

(4) identifying a Transformation Type for each of the identified independent variables and their related dependent variables;

(5) estimating a Coefficient for each of the identified independent variables and their related dependent variables;

(6) generating a Model Equation for each of the identified independent variables and their related dependent variables using the identified Transformation Type and estimated Coefficient;

(7) validating the generated Model Equation by applying it to the validation sample; and

(8) scoring customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing.

9. (ORIGINAL) The system of claim 8, wherein the Transformation Type is a mathematical operation that identifies an association between the identified related independent and dependent variables.

10. (ORIGINAL) The system of claim 8, wherein the Coefficient is a relative measure of the identified related independent and dependent variables' contributions to a likelihood of response.

11. (ORIGINAL) The system of claim 8, wherein the Coefficient's sign indicates whether the independent variable is positively or negatively correlated with the dependent variable.

12. (PREVIOUSLY PRESENTED) The system of claim 8, wherein the Model Equation is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders.

13. (ORIGINAL) The system of claim 8, wherein the logic for validating (7) further comprises logic for applying the generated Model Equation to the validation sample in order to predict a likelihood of response as compared to an actual response in the validation sample.

14. (ORIGINAL) The system of claim 8, wherein the logic for scoring (8) further comprises logic for applying the validated Model Equation to the customers retrieved from the database in order to predict responses from the customers in a future promotional campaign.

15. (PREVIOUSLY PRESENTED) An article of manufacture comprising a storage device embodying instructions that, when read and executed by a computer, result in the computer performing a method for creating a customer promotion response model for use in customer relationship marketing, comprising:

(a) generating, in a computer, an input data set for the response model, wherein the input data set is generated using an Analytic Data Set Template containing one or more Analytic Variables that include both primitives that are base variables and conditions that are predicates, aggregates or other functions, wherein the primitives and conditions determine how the Analytic Variables are derived from operational data to produce the input data set, and wherein the Analytic Variables are subdivided into independent variables and their related dependent variables;

(b) splitting, in the computer, the input data set into a test sample and a validation sample;

(c) identifying, in the computer, the independent variables and their related dependent variables using the test sample;

(d) identifying, in the computer, a Transformation Type for each of the identified independent variables and their related dependent variables;

(e) estimating, in the computer, a Coefficient for each of the identified independent variables and their related dependent variables;

(f) generating, in the computer, a Model Equation for each of the identified independent variables and their related dependent variables using the identified Transformation Type and estimated Coefficient;

(g) validating, in the computer, the generated Model Equation by applying it to the

validation sample; and

(h) scoring, in the computer, customers retrieved from a database using the validated Model Equation as a customer promotion response model for use in customer relationship marketing.

16. (ORIGINAL) The article of manufacture of claim 15, wherein the Transformation Type is a mathematical operation that identifies an association between the identified related independent and dependent variables.

17. (ORIGINAL) The article of manufacture of claim 15, wherein the Coefficient is a relative measure of the identified related independent and dependent variables' contributions to a likelihood of response.

18. (ORIGINAL) The article of manufacture of claim 15, wherein the Coefficient's sign indicates whether the independent variable is positively or negatively correlated with the dependent variable.

19. (PREVIOUSLY PRESENTED) The article of manufacture of claim 15, wherein the Model Equation is a mathematical representation of the association of the identified related independent and dependent variables that result in a statistical best fit of known responders versus non-responders.

20. (ORIGINAL) The article of manufacture of claim 15, wherein the validating step (g) further comprises applying the generated Model Equation to the validation sample in order to predict a likelihood of response as compared to an actual response in the validation sample.

21. (ORIGINAL) The article of manufacture of claim 15, wherein the scoring step (h) further comprises applying the validated Model Equation to the customers retrieved from the database in order to predict responses from the customers in a future promotional campaign.

EVIDENCE APPENDIX

None.

RELATED PROCEEDINGS APPENDIX

None.